COOLEY LLP
BOBBY GHAJAR (198719)
(bghajar@cooley.com)
TERESA MICHAUD (296329)
(tmichaud@cooley.com)
COLETTE GHAZARIAN (322235)
(cghazarian@cooley.com)
1333 2nd Street, Suite 400
Santa Monica, California 90401
Telephone: (310) 883-6400

MARK WEINSTEIN (193043)
(mweinstein@cooley.com)
KATHLEEN HARTNETT (314267)
(khartnett@cooley.com)
JUDD LAUTER (290945)
(jlauter@cooley.com)
ELIZABETH L. STAMESHKIN (260865)
(lstameshkin@cooley.com)
3175 Hanover Street
Palo Alto, CA 94304-1130
Telephone: (650) 843-5000

CLEARY GOTTLIEB STEEN & HAMILTON LLP
ANGELA L. DUNNING (212047)
(adunning@cgsh.com)
1841 Page Mill Road, Suite 250
Palo Alto, CA 94304
Telephone: (650) 815-4131

[Full Listing on Signature Page]

Counsel for Defendant Meta Platforms, Inc.

UNITED STATES DISTRICT COURT

NORTHERN DISTRICT OF CALIFORNIA

SAN FRANCISCO DIVISION

RICHARD KADREY, *et al.*,
 	
   Individual and Representative Plaintiffs,
 	
     v.
 	
 META PLATFORMS, INC., a Delaware corporation;
 	
           Defendant.
	Case No. 3:23-cv-03417-VC-TSH
	
	 DECLARATION OF CHAYA NAYAK IN SUPPORT OF META'S MOTION FOR PARTIAL SUMMARY JUDGMENT

I, Chaya Nayak, declare:

1.   I am over the age of 18 and am competent to make this declaration. I am a Director of Product Management in the Generative AI ("Gen AI") division of Meta Platforms, Inc. ("Meta"). I have been employed by Meta since October 2016. I have personal knowledge of the facts contained in this declaration in support of Defendant Meta Platform Inc.'s Motion for Partial Summary Judgment. I declare that the following is true to the best of my knowledge, information, and belief, and that if called upon to testify, I could and would testify to the following.

Professional Background

2.   I have a Master's degree in public policy & data science from University of California, Berkeley. For the past 10 years my professional work has largely focused on data science and technology product development.

3.   Before assuming my current role in 2023, I worked at Meta in various other capacities, including as Product Manager & Head of Open Research & Transparency (June 2018 to May 2023), and as a Product Manager for Gen AI focused on data and safety (May 2023 to December 2023).

4.   Gen AI was formed as part of a reorganization that occurred in 2023, shortly after the launch of Llama 1. Before Gen AI, much of our AI research was centralized within Meta's AI research lab, FAIR, which is focused on fundamental AI research. Gen AI was established to develop practical applications for AI, and assumed primary responsibility for development of Meta's Llama models upon its formation, though FAIR remains involved.

5.   As Director of Product Management, I have responsibility for overseeing product managers and managing engineers, data scientists, and other team members involved in the development of Meta's Llama models. Currently, I am responsible for overseeing development of Llama 4.

Background on Llama Development

6.   Llama 1 was released in February 2023. I was not involved in the development of Llama 1, but I am familiar with the model and the Llama 1 paper published by Meta upon its release titled, "LLaMA: Open and Efficient Foundation Language Model." A true and correct copy of the

1   Llama 1 paper is attached hereto as **Exhibit A**. Llama 1 was trained on 1.4 trillion tokens sourced

2   from a variety of publicly available datasets, which are detailed on page 2 of the paper along with

3   their respective proportions.

4   7.   As reflected in Exhibit A, 67% of Llama 1 training text was sourced from English

5   Common Crawl, with an additional 15% corresponding to the dataset C4, which is also compiled

6   from CommonCrawl data. Common Crawl is a 501(c)(3) nonprofit organization that crawls the

7   web collecting raw webpage data, metadata, and text extracts, and providing free access to the same

8   (https://commoncrawl.org/overview).

9   8.   As reflected in Exhibit A, 4.5% of Llama 1 training text was sourced from Github,

10   a leading cloud-based platform where coders store and share code, frequently on an open-source

11   basis (https://docs.github.com/en/get-started/start-your-journey/about-github-and-git).

12   9.   As reflected in Exhibit A, 4.5% of Llama 1 training data was sourced from

13   Wikipedia, a free online encyclopedia written by volunteers (www.Wikipedia.com).

14   10.   As reflected in Exhibit A, 4.5% of Llama 1 training data was sourced from a

15   combination of Gutenberg and Books3.

16   11.   As reflected in Exhibit A, 2.5% of Llama 1 training data was sourced from ArXiv,

17   a free online archive of math, science, and economics papers.

18   12.   As reflected in Exhibit A, 2% of Llama 1 training data was sourced from Stack

19   Exchange, a network of question-and-answer websites for sharing technical knowledge, geared

20   toward the programming community (https://stackexchange.com/about).

21   13.   Attached hereto as **Exhibit B** is a true and correct copy of a February 2023 blog post

22   published by Meta announcing the release of Llama 1, which describes Meta's belief that the open

23   release of Llama 1 would help democratize access to performant LLMs.

24   14.   When I joined Gen AI in 2023, Llama 2 development was underway and I assumed

25   significant responsibility helping to guide the model through to release, in particular with respect

26   to ensuring that the model adheres to Meta's safety standards. Meta invested heavily in finetuning

27   for Llama 2 in order to make the model safer and better suited for broad adoption.

28

1   15.   All of the Llama models released by Meta to date have been released to the open-
2 source community, which has allowed AI researchers and developers around the world to obtain
3 the model free of charge and to use and modify the Llama models for their own purposes, pursuant
4 to the terms of the applicable open-source license agreement. Llama models today can be
5 downloaded from a number of public sources including from Meta itself and from various third
6 party resources such as Hugging Face (a well-known website for AI researchers and
7 developers). To date, Llama models have been downloaded more than 1 billion times. A true and
8 correct copy of a recent press release announcing this milestone is attached hereto as **Exhibit C**.

9   16.   All of the Llama models released to date have been released to the open-source
10 community pursuant to open-source license agreements, the terms of which have varied from
11 version to version. The initial release (Llama 1) was released under a license that permitted
12 research (non-commercial) purposes only. Llama 2 was released under a different license that
13 allowed the model to be freely used for commercial purposes, subject to various conditions
14 including the requirement that the developer comply with applicable laws and regulations and
15 adhere to an Acceptable Use Policy (AUP) established by Meta that forbids the model from being
16 put to various prohibited uses such as violating other people's rights or being used to engage in or
17 facilitate unlawful conduct. True and correct copies of the Llama 1 and Llama 2 licenses are
18 attached hereto as, respectively, **Exhibits D** and **E**. A copy of the Meta AUP, referenced in the
19 Llama 2 license, is attached as **Exhibit F**.

20   17.   In conjunction with the release of Llama 2, Meta published a paper titled "Llama 2:
21 Open Foundation and Fine-Tuned Chat Models," a true and correct copy of which is attached hereto
22 as **Exhibit G**, which details the development process and features of the model.

23   18.   A preview version of Llama 3 was released in April 2024. Llama 3.1 was released
24 in August 2024 alongside the paper "The Llama 3 Herd of Models," a true and correct copy of
25 which is attached hereto as **Exhibit H**.

26   19.   It was shortly followed by the release of Llama 3.2 in September 2024. A true and
27 correct copy of Meta's press release announcing the release of Llama 3.2 is attached hereto as
28 **Exhibit I**. The Llama 3.2 release included small and medium-sized multimodal LLMs capable of

analyzing both text and images. It also included much smaller, text-only models (1B and 3B) that are capable of being stored and run on mobile devices.

20.   The most recent model, Llama 3.3, was released in December 2024. Llama 3.3 is a text-only model available in only one size, 70B, and provides enhanced performance relative to prior versions of Llama 3 that are similar in size.

21.   The Llama 3 models, like Llama 2, were released on an open source basis that permits commercial and non-commercial uses, subject to the terms of the applicable license, which also requires adherence to the Acceptable Use Policy (AUP) mentioned above. A true and correct copy of the Llama 3 license is attached hereto as **Exhibit J**.

22.   Each version of Llama that Meta has seen significant improvements in performance over previous Llama versions (as measured across a variety of industry benchmarks). The amount of training data used to train Llama models has also increased from version-to-version, with Llama 1 trained on 1 trillion tokens, Llama 2 on 2 trillion tokens, and Llama 3 on 15 trillion tokens. Meta has also made available variations in model sizes to provide developers and researchers with more flexibility in utilizing the Llama models and incorporating them into their own projects and implementations. For example, the Llama 3 model family is available in a number of different sizes up to 405 billion parameters, as well as smaller models that can be run on mobile devices.

23.   Meta has also released "Meta AI," an AI assistant that is available to people who are not AI researchers and developers, that allows users to enter prompts and obtain responses. Meta AI is currently available through a web-based interface as well as through various Meta products including Facebook, Instagram, and WhatsApp.

Third-Party Innovations Using Llama

24.   Meta has a long history of contributing to open source technology development. For example, PyTorch is a Meta-developed open-source machine learning library that has become widely adopted by the AI research and development community. Meta's open release of the Llama models continues that tradition, with the objective of fostering innovation in the development and deployment of AI.

25.   Countless organizations and businesses rely on Llama models to conduct research and develop their own products.

26.   For example, the communications platform Zoom utilizes Llama to power its generative AI assistant which, among other things, enables users to create summaries of meetings. A true and correct copy of a press release detailing Meta's collaboration with Zoom is attached hereto as **Exhibit K**.

27.   Numerous companies, including DoorDash and AT&T, have adapted Llama to improve productivity and customer service, while The Washington Post has used Llama to create "Ask the Post" (https://www.washingtonpost.com/ask-the-post-ai/), a chatbot that answers questions based on information from The Washington Post's article archives dating back to 2016. True and correct copies of Meta blog posts detailing some of these projects are attached hereto as **Exhibit L**.

28.   Yale School of Medicine partnered with EPFL's School of Computer and Communication Sciences to develop Meditron, an open source LLM built on Meta's Llama models, to improve access to evidence-based information for clinical decision making. A true and correct copy of a Meta blog post detailing this project is attached hereto as **Exhibit M**.

29.   As another example, the company Envision, a pioneering assistive learning technology company, is integrating Llama into wearable devices that improve accessibility for the blind by describing visual information. A true and correct copy of the landing page of the Envision website located at www.letsenvision.com is attached hereto as **Exhibit N**.

30.   Other examples of innovative uses of Llama are featured on the Meta webpage located at https://www.llama.com/community-stories/, a true and correct copy of which is attached hereto as **Exhibit O**.
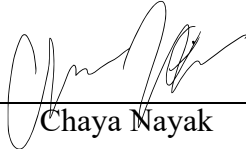
31.   As a part of Meta's commitment to supporting open source development, the company established the Llama Impact Grant Program in October 2023. Pursuant to this initiative, Meta solicited proposals from organizations to use Llama to address real world problems, with winning proposals receiving grants of up to $500,000. True and correct screenshots of the Meta

website detailing the Llama Impact Grant Program, its participants, and grant winners, are attached hereto as **Exhibit P**.

32.   As described in Exhibit P, recipients of Meta's inaugural impact grant award included the following:

- Wadhwani AI for its AI-enabled reading assessment to enhance English fluency and comprehension skills in public schools in Gujarat, India. The teaching module will be designed to utilize Llama to generate level-appropriate paragraphs for students, allowing them to focus on and gain mastery of words that they find challenging.

- Digital Green for its Farmer Chat project, which aims to develop a multilingual AI chatbot that provides customized, on-demand agricultural advisory services to small-scale farmers in Sub-Saharan Africa. Llama will be fine-tuned on a vast repository of agricultural data to adapt the model to the specific needs and contexts of different regions. Llama's conversational capabilities will then be used to provide comprehensive and practical responses in a variety of languages including Hindi, Swahili, and Kikuyu.

- The Dana-Farber Cancer Institute (DFCI) for its proposal to develop a new open source AI platform to computationally match patients with cancer to clinical trials. They plan to use Llama 3 to summarize and comprehend unstructured clinical notes and unstructured clinical trial eligibility criteria, enabling rapid identification of appropriate trial options for individual patients.

- Jacaranda Health for an expansion of its digital health tool, called PROMPTS, which uses SMS behavioral nudges and an AI-enabled clinical helpdesk to empower new and expecting mothers across Kenya, Ghana, and Eswatini to seek and connect with the right care at the right time.

1

2

3   I declare under penalty of perjury that the foregoing is true and correct. Executed on this 20

4   day of March 2025, at 10:00 PM.

5   _____
                                                                                          Chaya Nayak
6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

ATTORNEYS AT LAW

                                                  7   DECLARATION OF CHAYA NAYAK
                                                                            CASE NO. 3:23-CV-03417-VC-TSH